A Network-based Distributed, Media-rich Computing and Information Environment

Richard L. Phillips
Los Alamos National Laboratory
Los Alamos, NM USA

OVERVIEW

Sunrise is a Los Alamos National Laboratory (LANL) project started in October 1993. It is intended to be a prototype National Information Infrastructure development project. A main focus of Sunrise is to tie together enabling technologies (networking, object-oriented distributed computing, graphical interfaces, security, multi-media technologies, and data-mining technologies) with several specific applications. A diverse set of application areas was chosen to ensure that the solutions developed in the project are as generic as possible. Some of the application areas are materials modeling, medical records and image analysis, transportation simulations, and K-12 education. This paper provides a description of Sunrise and a view of the architecture and objectives of this evolving project.

The primary objectives of Sunrise are three-fold:

- To develop common information-enabling tools for advanced scientific research and its applications to industry.
- To enhance the capabilities of important research programs at the Laboratory
- To define a new way of collaboration between computer science and industrially-relevant research.

The basic paradigm being developed involves a document-centric user interface which allows arbitrary object support including embedded applications, multimedia video/voice fragments and links to a wide information space. We will eventually develop an information kiosk based on an ATM network so that all the participants can exchange, publish, or interact with applications and data. This will function on heterogeneous platforms, provide for constrained access to data through security mechanisms, and be extensible. The data-mining technology will include the ability to quickly browse large complex image databases with various feature extraction capabilities, provide advanced, selective, compression algorithms, and the ability to merge and purge large complex datasets. The focus is on real scale problems, reflecting our belief that only experience with real problems encompassing huge data sets will facilitate true progress. Therefore, our focus will be on developing prototype tools that can be tested in real-sized testbeds. We believe the benefits to the Laboratory as a whole will be in helping to define the way research will be done in the future, in enhancing the competitiveness of Laboratory research programs, and in enhancing the nation's ability to use advanced information technologies for applications of importance to industry. The project is organized into subgroups consisting of people working on the enabling technologies and people working on each of the several representative application areas. Figure 1 illustrates some of these activities and relationships.

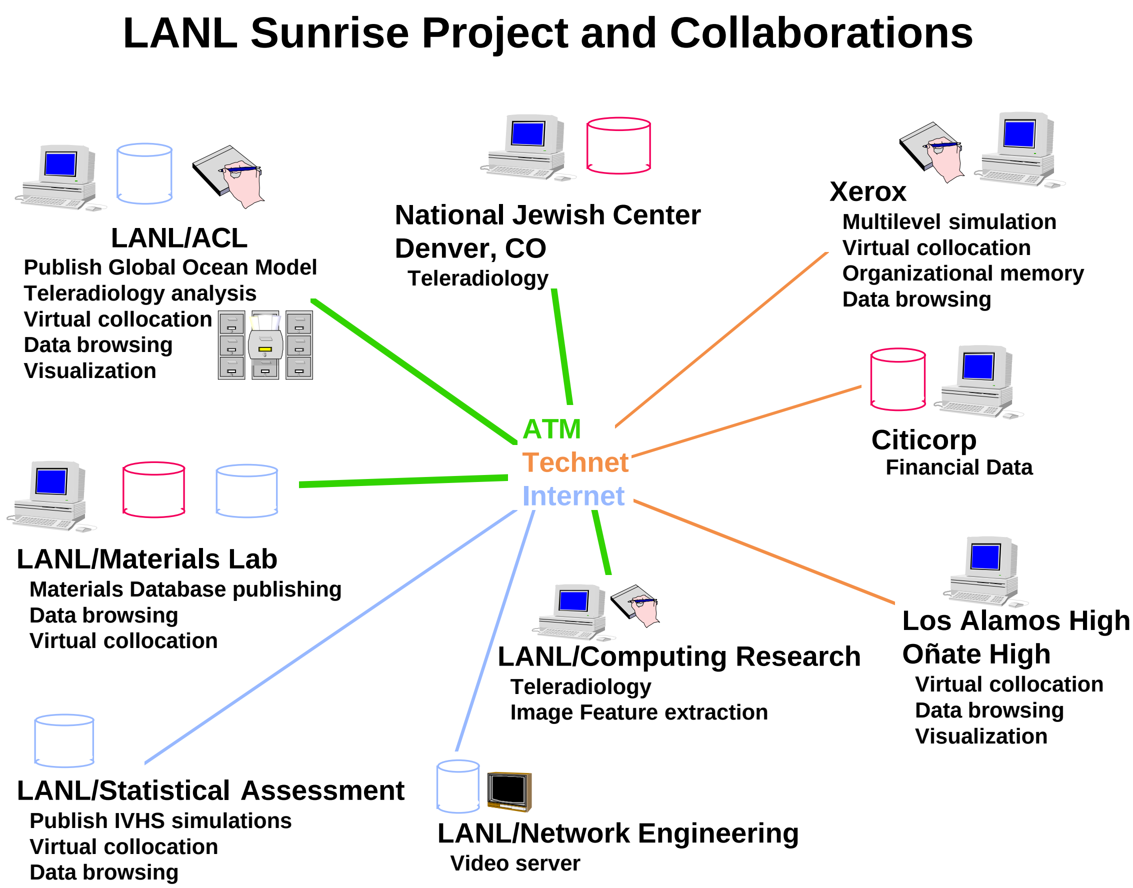

Figure 1. Sunrise Components

SUNRISE STRUCTURE

The Sunrise project is organized according to an integrated, layered approach. The three basic layers are:

- hardware, e.g. networking fabric and computer systems.
- services, e.g. security and database systems.
- applications, the reason for it all, e.g. manufacturing, medicine.

This approach to structuring the project is illustrated in Figure 2. Some aspects of each layer will be discussed briefly.

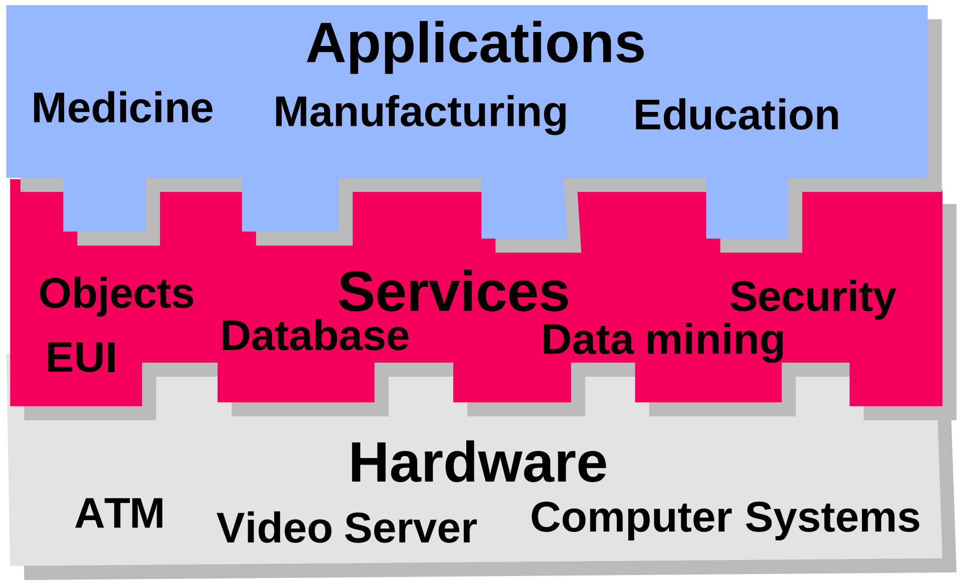

Figure 2. Sunrise Layers of Technology

Hardware — Computing

Sunrise is making use of several high performance computing resources as network-accessible computing servers. The most significant of these are a Thinking Machines CM-5 and a Cray Research T3D. The CM-5 is a 1056 node machine with 33 GB of memory and 440 GB of disk (accessible at up to 200 MB/s). It has 4 High Performance Parallel Interface (HIPPI) 800 Mb/s channels for sending data to other devices. It has a peak speed of 132 Gflops and has sustained over 60 Gflops on real applications. It is being used for global climate modeling, materials modeling, gas and oil seismic and reservoir modeling, as well as distributed visualization and medical image analysis.

The T3D is a 128 node machine with 128 DEC Alpha processors and 8 GB of memory. It has 90 GB of attached RAID disk and one HIPPI channel for high speed communication. It is primarily being used for medical image analysis and various industrial applications.

Hardware — Networking

The Sunrise project will involve the use of several technologies that will enhance the effectiveness of computers in the office environment of a laboratory staff member. A critical part of this effectiveness is communication and information retrieval. Using email and other Internet applications of today as models, we hope to expand the capabilities of a workstation and its network to provide such things as video on demand, teleconferencing, visualization postprocessing, and other tools that free up time for the user to be more effective in other endeavors.

At the present time a high performance network fabric is not ubiquitous, so not all of the services just enumerated can be provided to all clients. Therefore, Sunrise currently uses whatever connectivity is available and scales services accordingly. Presently we are using local area 10 MB/s Ethernet and 100 Mb/s FDDI, and wide area 1.5 Mb/s (DS1) and 45 Mb/s (DS3) Internet services, and 34 Mb/s SMDS services. A network technology that, hopefully, will support the full service vision is one where various data rates and qualities of service are supported. The emerging Asynchronous Transfer Mode (ATM) appears to provide that support and will eventually be used throughout the Sunrise environment.

Initially, the ATM switches will only support variable bit rate (VBR) service and will be operating at 100 Mb/s. The constant bit rate (CBR) service for support of voice and video will come later as will higher data rates of 155 Mb/s under the Synchronous Optical NETwork specifications (SONET). The SONET aspect is an important issue because the telephone companies from the local carriers (US West) to the long haul carriers (AT&T, MCI, Sprint, etc) are now building a SONET infrastructure to support wide area networks using ATM. Their goal is to have the ATM switch located at the central office of the telephone company rather in the computer room. The important point, however, is that the network technology will be the same for local area networks (LANs) as well as wide area networks (WANs), which is definitely not the case today. Using similar technologies in the WAN and LAN make interconnectivity seamless and much more robust

Services — Distributed Objects

The Sunrise project is based on the premise that a distributed object approach will be necessary for a National-scale information system. We are trying to anticipate the future and are using the Object Management Group's (OMG) evolving Common Object Request Broker Architecture (CORBA)[1] since it is the emerging industry standard in this area. We are also investigating the use of the OpenDoc[2] compound document standard. The OpenDoc storage model is based on Bento. Both the OpenDoc and Bento efforts are being coordinated by the Component Integration Lab (CIL).

The usefulness of the NII will be greatly enhanced when users of applications can quickly and easily locate and access data and computing resources in the network. Those resources may be, for example, public or private electronic libraries, data conversion services, or various other application specific services. Since thousands of services built by hundreds of organizations will exist on the network it is important to facilitate ways of building applications by using software components built by different organizations.

We must build an infrastructure that allows systems to be built and modified with interchangeable components in much the same way that we combine stereo components to build an audio system. When building an audio system we can interchangeably add or subtract components without redesigning the system or even replace a tape player with a CD player or some other device without changing anything else in the system.

Users accessing the NII will be combining desktop machines such as PCs and workstations with various data and computational servers into a widely distributed yet integrated system. Components of the system will often have different CPU architectures, different data communication infrastructures and data rate capabilities, different operating systems, and be widely distributed. Nevertheless, users' applications will need to be able to call on those resources as if the resources were locally available.

This composable component-based system requires further evolution of distributed computing environments. To accomplish this evolution, we are basing the Sunrise environment on the notion of distributed objects. Properly designed object-based systems have the unique advantage that changes in one area of the system do not affect other areas. A properly designed set of object classes will enable us to link resources in a modular way to manage complexity and yet allow great extensibility. We are building the Sunrise distributed computing infrastructure on Orbix[3], Iona Technologies' CORBA implementation.

Early emphasis in the development of objects has been on the Sunrise *TeleMed* project with the National Jewish Center in Denver. The system allows CAT scan and x-ray images of lungs to be submitted from anywhere in the network to an electronic repository. The image is then be processed to compute a "signature" that can then be used to find other similar images.

Services — Executive User Interface

The overall goal of this activity is to develop a general, media-rich user interface system which has capabilities that meet the needs of all Sunrise application clients. In general, this is an executive user interface (EUI) , one that has a set of convenience capabilities common to all application areas, while still providing application clients easy access to their preferred user interfaces. One of the most important convenience capabilities is to facilitate telecollaboration between users. In its simplest form this feature permits peer-to-peer collaboration in a common screen space. More advanced capabilities will include many-to-many collaboration and video teleconferencing. The Sybase product Gain*Momentum*[5] is being used for interface development.

Among the features of the EUI are:

- Compound (media-rich) document executive; provides a complete multimedia authoring and display capability. Data stored in Bento (the OpenDoc persistent storage system) documents can be dynamically assembled and viewed.
- A logbook facility, which draws upon the capabilities described above, will allow a researcher to maintain a time-stamped multimedia record of all activities. Since the logbook will be retained as a Bento file, it can be accessed by other applications
- Generic Sunrise data navigator; provides a search and retrieval tool, including mixed media search capability. An appropriate viewer is invoked depending on media type.
- Customized, application-specific user interfaces can be defined and easily incorporated. Launch buttons can be added as needed to the basic EUI.
- Other EUI convenience features are a MIME-compliant mailer and viewer, multimedia annotations; post-its, sketchers, audio, video, animation, etc. For intra- and inter-document navigation, thumbnails, hyperlinks, and World Wide Web links are provided.

An example of the appearance of an EUI and the possible contents of a logbook is shown in Figure 3. Here the researcher elected to log a series of radiographic studies, each one represented by a thumbnail icon that points to the study. A Post-it note was added to the page by dragging an instance of the note icon from the right of the tools bar. The Post-it is shown in its open state to reveal its contents.

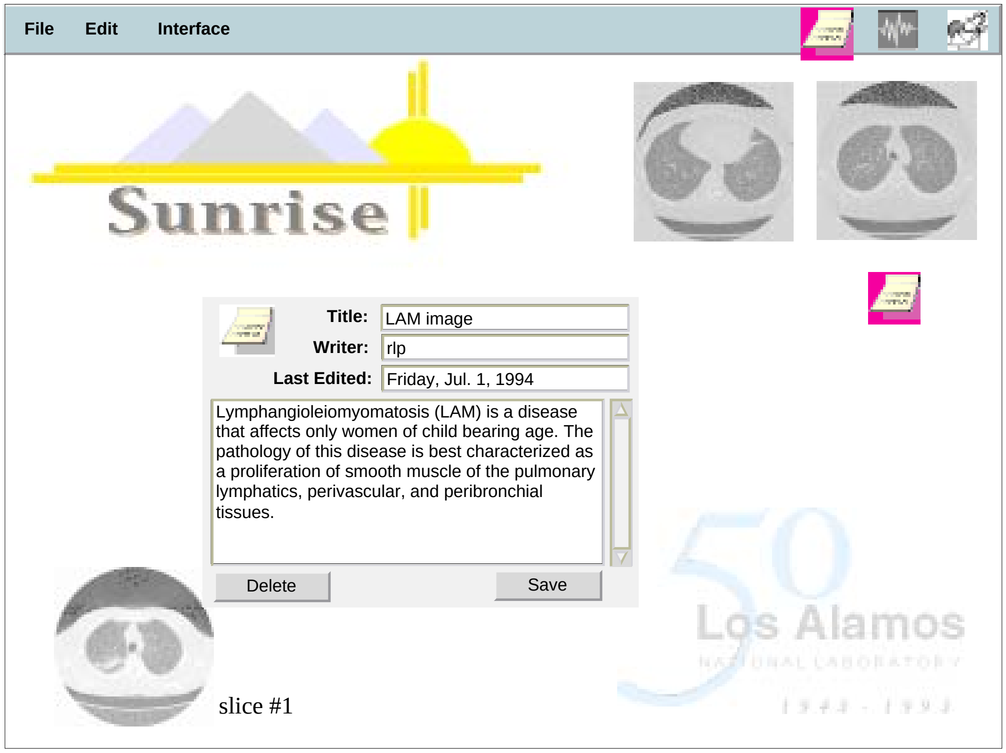

Figure 3. Sunrise Executive User Interface

Services — Security

For the National Information Infrastructure to truly be a valuable national resource, it must support a diverse set of users and usages. The network will therefore have traffic ranging in sensitivity from public information (such as income tax forms) to sensitive medical and financial records. The loss or corruption of either would cause serious (and in some cases irreparable) damage to individuals, companies and government. Computer security is therefore essential to the success of the NII.

The goal of the computer security part of the Sunrise project is to provide host and network security for services and for machines offering those services. To that end, one of the primary projects is to provide the tools to build secure applications easily and and in a fool-proof way.

We have adopted an ANSI X9-based certifying authority system for secure public key distribution and a Kerberos v4 and v5-based library for securely communicating with the certifying authority. All security features are implemented as a service in the Orbix system (the CORBA implementation mentioned earlier) and are thus available for security enforcement among all objects. In addition, we provide a library for secure communication (includes DES and RSA encryption routines) as well as connect and accept routines for automatically transferring credentials and session keys securely.

Applications — Telemedicine

The most well-developed application in the Sunrise suite is a medical application called *TeleMed*. This application serves as an excellent case study of all the Sunrise technologies previously described and thus warrants description in some detail. *TeleMed* grew out of a relationship with physicians at the

National Jewish Center for Immunology and Respiratory Medicine (NJC) in Denver, Colorado. These physicians are experts in pulmonary diseases and radiology, helping patients combat the effects of tuberculosis and other lung diseases throughout the Nation. These individuals are an expensive and scarce resource, who often travel around the country to share their expertise with other physicians.

To make the knowledge and experience at the National Jewish Center available to a wider audience, Los Alamos National Laboratory is developing a telemedicine system called *TeleMed* which is based on a national radiographic repository located at Los Alamos . Without leaving their offices, participating doctors can view radiographic data via a sophisticated multimedia interface. For example, a doctor can match a patient's radiographic information with the data in the repository, review treatment history and success, and then determine the best treatment. Furthermore, the features of *TeleMed* that make it attractive to clinicians and diagnosticians make it valuable for teaching and presentation as well. Thus, a resident can use *TeleMed* for self-training in diagnostic techniques and a physician can use it to explain to a patient the course of their illness.

LANL's expertise in the areas of high-performance computing, advanced networking, image processing, and multimedia information systems has proved to be an important asset for *TeleMed*. These technologies come into play in the following ways:

- **image processing** - As databases are developed to store large volumes of information, new methods for mining huge amounts of data will become increasingly important. Tools will be needed for analyzing text, imagery, 1-D signals, and other numerical information (LDL & HDL cholesterol counts, blood pressure, body temperature, etc.) One such key tool that is being used in the *TeleMed* project is automated data analysis of radiographic images. This tool provides the physician with an agent that browses radiographic data and calls the physician's attention to suspect areas. One typical agent can analyze CT images from patients with lymphangioleiomyomatosis (LAM) disease. Pulmonary CT studies reveal that the primary signature of LAM is the presence of many cysts throughout the lung. Members of the Sunrise team have created algorithms that automatically locate the cysts in the pulmonary CT data and record their size.[4] One can compute a histogram of the cyst sizes, which is called a cystogram. The cystogram provides a meaningful quantitative measure of the progress of the disease.

- **high-performance computing** - Especially important to the *TeleMed* project is the fact that LANL houses the massively parallel computing systems mentioned earlier, the Thinking Machines CM-5 and the Cray T3D. Since medical image analysis techniques (such as the LAM agent described above) can be computationally intensive, Sunrise acts as a server of massively parallel computing cycles to physicians in the *TeleMed* network. Access to such a compute server allows physicians to do near real time consultation with image analysis agents.

- **advanced networking** - Over the last few years LANL has made numerous contributions to high performance networking technology. The importance of this technology to the *TeleMed* project is clear, since multi-megabyte radiographic data files must be transferred rapidly to the diagnostician. Moreover, as noted earlier, Sunrise is aggressively exploring the role of ATM communications technology for use both in image data transfer and desktop video telecollaboration.

- **multimedia information systems** - The July 1994 issue of R & D magazine[6] features, as a cover story, LANL-developed MediaView[7], a comprehensive multimedia document authoring and management system. MediaView represents the advanced multimedia expertise Los

Alamos is applying to the *TeleMed* project. Through MediaView, the concept of a document-centric user interface was independently developed. This type of interface is used in the *TeleMed* project and affords users access to a complete range of media-rich components. Physicians can, for example, affix digital Post-it notes to a patient record or x-ray. Voice annotations can be dictated and directly attached to a patient document. Also, a physician can digitally sketch on an x-ray, pointing out important features to a collaborator or patient.

Some of the capabilities of *TeleMed* can be illustrated by looking at a series of user interface components that are available to the user. The user begins a *TeleMed* session by selecting a database site from the interface shown in Figure 4. This sets in motion an Orbix-based transaction for vending all patient record objects from the selected site to the requesting client, shown listed in Figure 4 as Patient #1, Patient #2, etc.

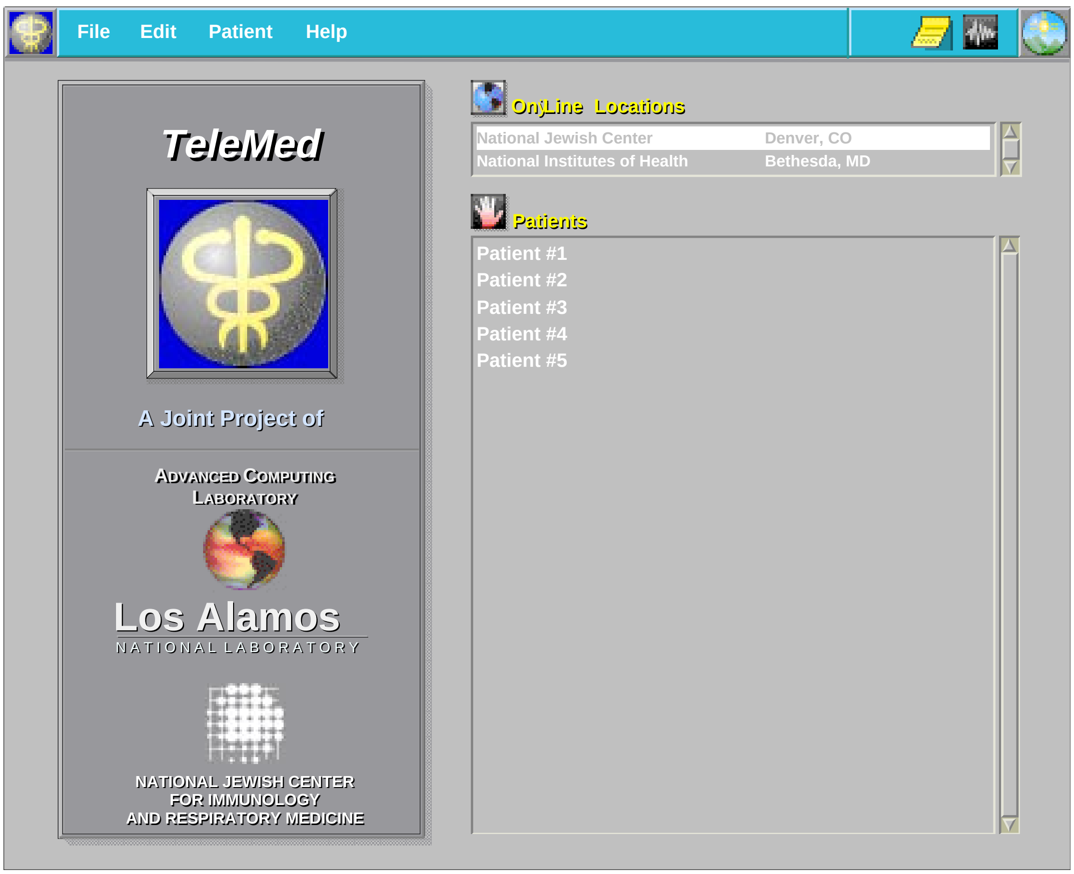

Figure 4 Initial *TeleMed* Interface

To understand the coordination of distributed object activities with user interface activities we can consult Figure 5, a graphical representation of *TeleMed* objects. In this diagram the arrows represent an inheritance relationship and the other lines represent a reference or containment. Textual data from the PatientInfo object, i.e., the patient’s name, was retrieved and used to populate the Patient list in Figure 4.

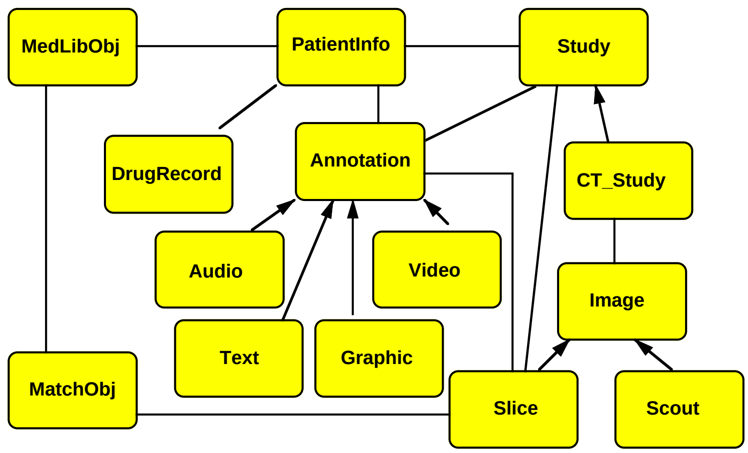

Figure 5 *TeleMed* Objects

A patient's treatment record appears by double-clicking on the patient's name in the interface in Figure 4. The user interface manifestation of the DrugRecord object in Figure 5 is called a Graphical Patient Record (GPR) and is shown in Figure 6.

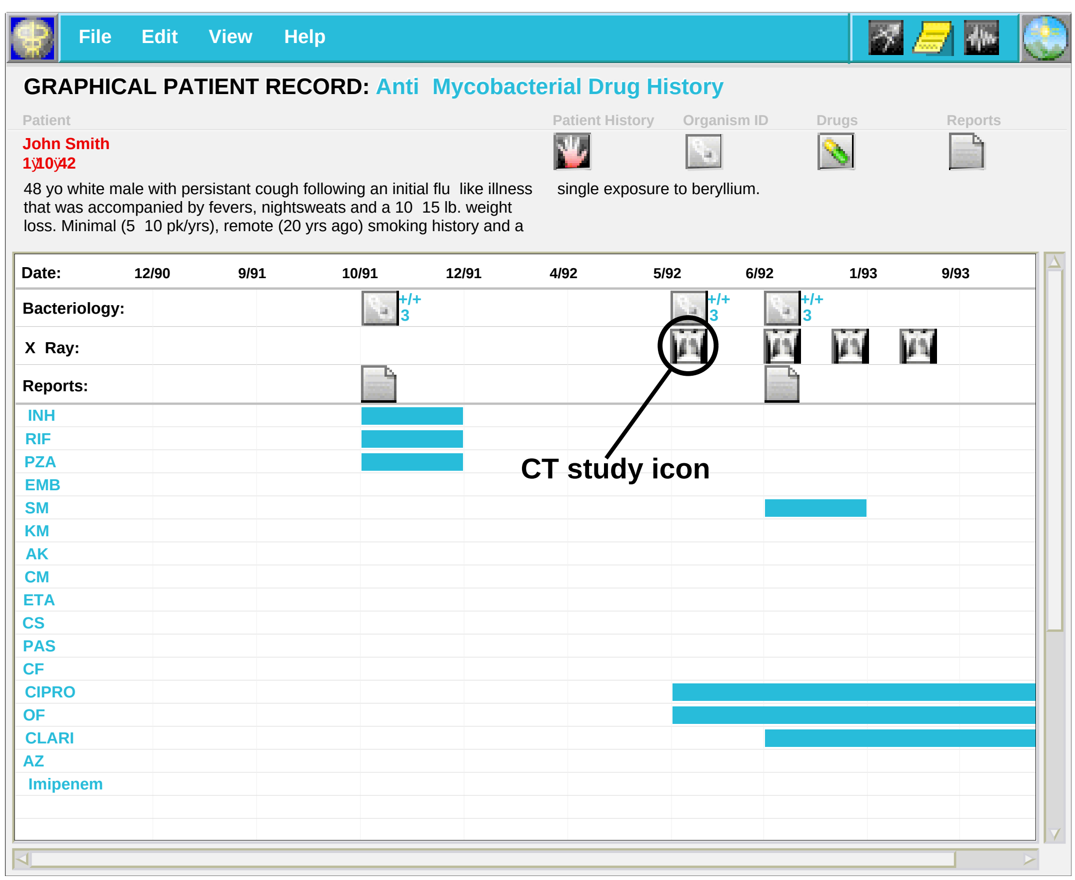

Figure 6 *TeleMed* Graphical Patient Record

The GPR is an excellent example of Sunrise/*TeleMed* media-rich document and distributed object technology. The GPR is a virtual document, a patient record that is empty until it is dynamically populated by requests for distributed objects. The DrugRecord object in Figure 5 contains the information necessary for “harvesting” this data from all appropriate sites. Thus, laboratory reports may be retrieved from the National Institutes of Health near Washington, DC while radiographic data may reside at the NJC in Denver. So, for example, when all patient data are retrieved, icons representing laboratory tests, radiographic studies, drug treatments, etc. are drawn on the GPR template. Each of these icons is mouse-sensitive and, when clicked, call up additional user interfaces and related data.

Before looking at these interfaces it will be helpful to know more about what goes on at the distributed objects level. In Figure 7 we show the relationships between the client process (TeleMed GUI) and the two controlling objects, MedLibObj and MatchObj. Any of these three entities can reside at any location. In fact, the TeleMed GUI can communicate with any number of MedLibObj objects, which, in turn, can call upon the services of any number of MatchObj objects. Suppose, now, the user clicks on a CT study icon in the GPR in Figure 6. This causes a request to be sent to the current MedLibObj to retrieve that patient’s CT study from the corresponding persistent object store. That transaction causes the user interface shown in Figure 8 to appear.

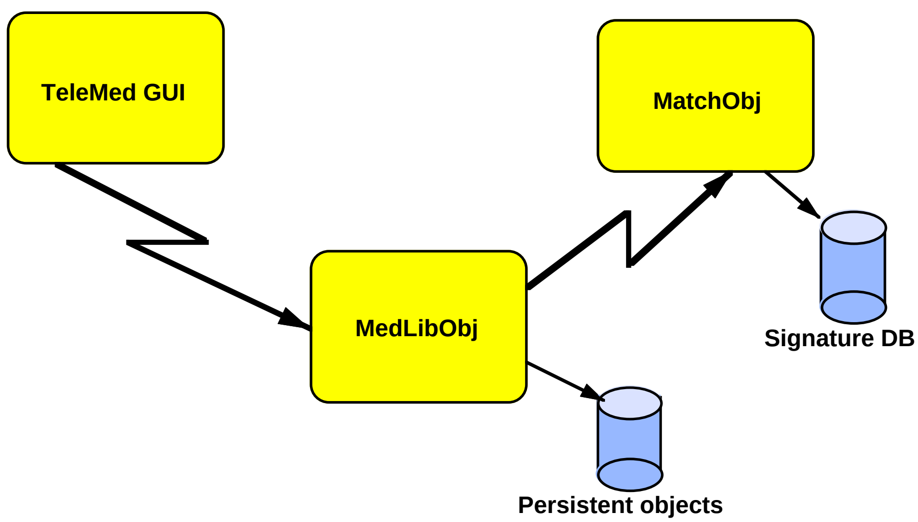

Figure 7 ***TeleMed*** **Object Distribution**

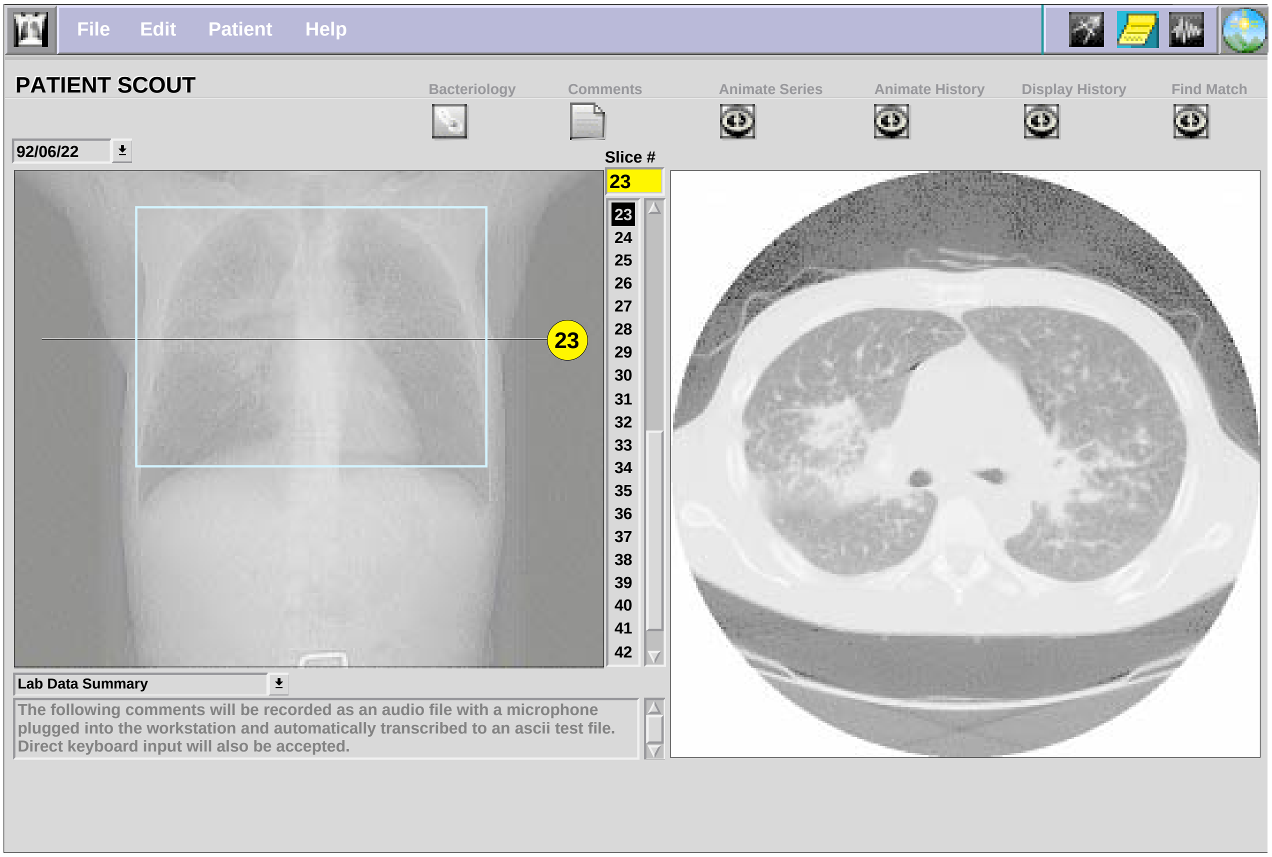

Figure 8 ***TeleMed*** **CT Study Interface**

The image on the left of Figure 8 is a *scout*, so named because it was originally used by the CT technician as a guide in determining where to produce full transverse slice images of the patient. In this interface the scout is similarly used, but now as a guide for the physician in selecting slices to view from the database. To do this, the horizontal cursor is dragged up or down to the desired location and released. Here, slice number 23 was selected and is shown on the right.

We conclude our discussion of the *TeleMed* application by describing one of its most powerful features. This feature allows a user to perform a “query by example” search of an image database. Many of the Sunrise enabling technologies mentioned earlier are represented in this feature — massively parallel computation servers, image analysis agents, and distributed object computing. To be specific, the MatchObj object shown in Figure 7 encapsulates the image analysis agent as a member function. For best performance MatchObj will typically reside on a massively parallel computer because the matching algorithm is inherently parallel. The signature database, which contains representative features of each image, usually resides on the same machine as MatchObj. Finally, the user invokes this entire matching operation simply by clicking the “Find Match” button in the upper right of Figure 8. The selected slice is used as the query image. The result of a matching operation is shown in Figure 9.

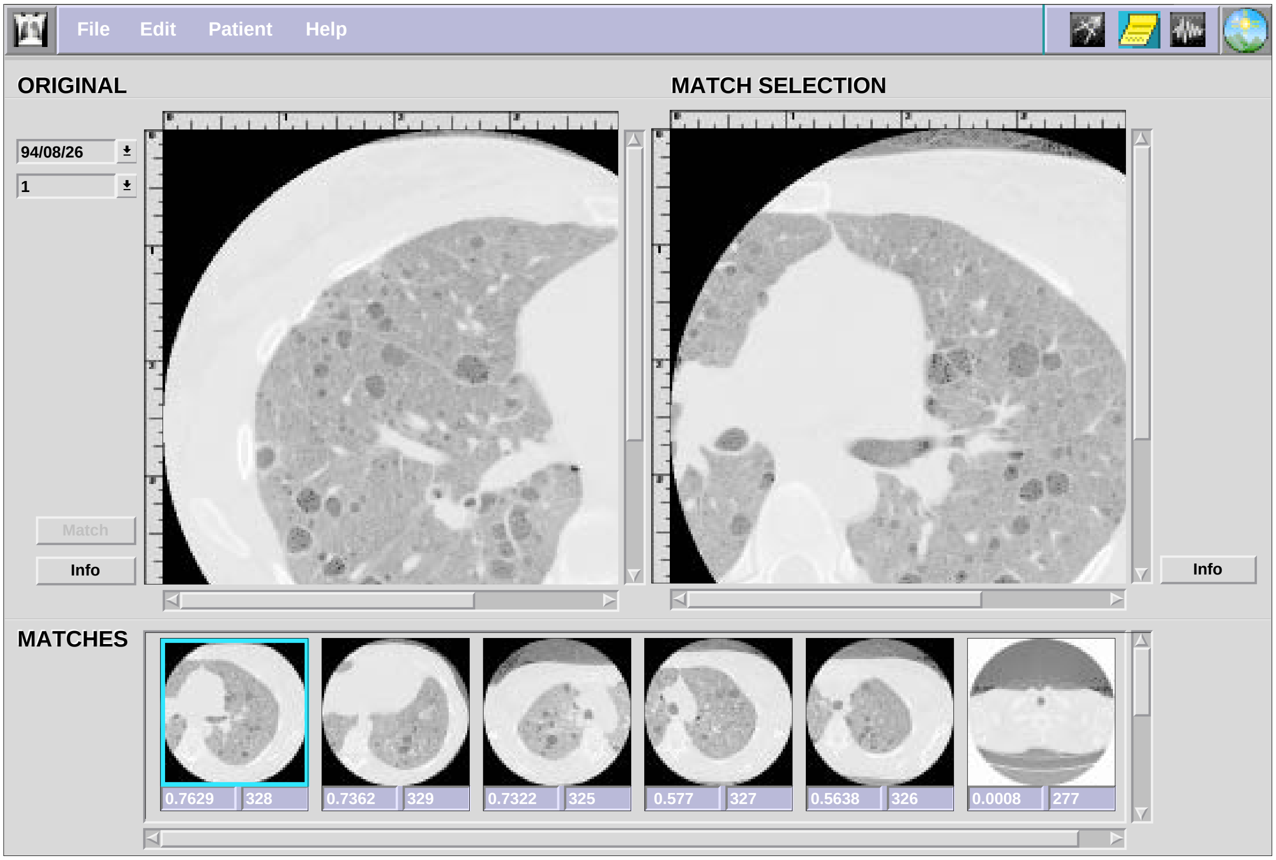

Figure 9 ***TeleMed*** **Image Matching Interface**

In Figure 9 the upper left image is the same one the user specified as the example query image. The result of the match is summarized by the thumbnails in the lower scrolling window. Clicking in a selected thumbnail causes its full-size representation to appear in the upper right comparative inspection window.

TeleMed has recently been deployed via DS1 service between Los Alamos and the National Jewish Center. At this point, physicians at National Jewish are learning the user interface and reacting to its design and shortfalls. As the *TeleMed* system reaches a productive state, other sites will be added to the *TeleMed* network. These will all be sites who are treating and doing research on multi-drug resistant tuberculosis (MDR-TB). Those intending to join are the National Institutes of Health in Bethesda, MD, the Center for Disease Control in Atlanta, GA, the Bureau of Tuberculosis Control in New York, NY, and the Department of Health Services in Los Angeles, CA. With these sites in place, physicians will truly be able to telecollaborate on MDR-TB cases, using the Los Alamos radiographic repository as a common point of contact.

Applications — Others

Other Sunrise applications include:

- materials modeling – tools are being developed to rapidly analyze material properties obtained from simulations and instruments such as a tunnelling electron microscope. Accomplishments include the development of techniques to correctly identify fractures in materials, even in the presence of poor quality data.
- K-12 education – virtual laboratories have been developed to demonstrate multimedia databases in a real world context. This application currently uses the Internet to deliver real-time digital video and synchronized sound to high schools in New Mexico and Hawaii.
- electronic documents – HyperTeX software is being used for automatically creating hypertext documents from existing TeX and LaTeX documents. Software has been developed to convert these hypertext documents to Adobe's PDF format. Eventually, WWW URLs will be incorporated as well.

TO LEARN MORE

The Sunrise project is discussed, with associated graphics, in a Mosaic server located at Los Alamos. The URL for Sunrise is **http://www.acl.lanl.gov/sunrise**.

The URL for the *TeleMed* application is **http://www.acl.lanl.gov/sunrise/Medical/telemed.html**.

ACKNOWLEDGEMENTS

The work described here represents the contributions of many people, indeed, the entire Sunrise team. While the writer is among those contributors, his role here has been largely that of rapporteur. Others who must be mentioned are Dave Forslund, Bob Tomlinson, Pat Kelly, Allen McPherson, Tim Merrigan, Pat Eker, Dave Kilman, Paul Hinker, John Reynders, Jon Bradley, Jonathan Greenfield, Steve Tenbrink, Niels Jensen, Paul Ginsparg, Mark Doyle, and others.

REFERENCES

1. Object Management Group, Ed.. *Common Object Request Broker Architecture*, Wiley-QED, 1991.
2. Anon., "OpenDoc Overview", Component Integration Laboratories, San Francisco, CA, 1994.
3. Horn, Chris, "Focus on Open Systems", *Object Magazine*, Sept-Oct, 1993. Also see http://www.iona.ie/www/Orbix/arch/Summary.html.
4. Everhart, J., Cannon, M., Newell, J., and Lynch, D., "Image Segmentation Applied to CT Examination of Lymphangioleiomyomatosis (LAM)", to be published.
5. Anon., *GainMomentum User's Guide*, Sybase, Inc. Emeryville, CA, 1994.
6. *R & D*, Cahners Publishing, July, 1994.
7. Phillips, R., "MediaView: A General Multimedia Digital Publication System", *Comm. ACM*, Vol. 34, No. 7, July, 1991.